

Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”

ОСАУЛЕНКО Вячеслав Миколайович

УДК 004.8

**МОДЕЛІ БІОЛОГІЧНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ
ПРОСТОРОВО-ЧАСОВОЇ АСОЦІАТИВНОЇ ПАМ'ЯТІ**

05.13.23 – системи та засоби штучного інтелекту

АВТОРЕФЕРАТ
дисертації на здобуття наукового ступеня
кандидата технічних наук

Київ – 2019

Дисертація є рукопис

Робота виконана в Навчально-науковому комплексі «Інститут прикладного системного аналізу» Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Науковий керівник: доктор фізико-математичних наук, професор,
Макаренко Олександр Сергійович,
Навчально-науковий комплекс «Інститут прикладного системного аналізу» Національного технічного університету України «Київський Політехнічний Інститут імені Ігоря Сікорського»,
зав. відділу прикладного нелінійного аналізу

Офіційні опоненти: доктор технічних наук, професор
Корабльов Микола Михайлович,
Харківський національний університет радіоелектроніки
професор кафедри комп'ютерних інтелектуальних технологій;

кандидат технічних наук,
Андрійчук Олег Валентинович,
Інститут проблем реєстрації інформації Національної академії наук України,
старший науковий співробітник лабораторії систем підтримки прийняття рішень відділу спеціалізованих засобів моделювання.

Захист відбудеться 11 грудня 2019 о 15⁰⁰ годині на засіданні спеціалізованої вченої ради Д 26.002.03 при Національному технічному університеті України “Київський політехнічний інститут імені Ігоря Сікорського” за адресою: 03056, м. Київ - 56, пр. Перемоги, 37, корп. №35, ауд. 001

З дисертацією можна ознайомитися у бібліотеці Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського” за адресою: 03056, м. Київ - 56, пр. Перемоги, 37.

Автореферат розіслано 08 листопада 2019 р.

Вчений секретар
спеціалізованої вченої ради Д 26.002.03
д.ф.-м.н., професор

В.О. Капустян

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. За останні десятиліття кількість інформації значно збільшилися завдяки розвитку мережі інтернет та здешевленню засобів персональної генерації та поширення контенту. Тому значну увагу приділено алгоритмам, що обробляють та структурують великі масиви даних. Сюди включають алгоритми розпізнавання, кластеризації та, поки в зародковому стані, алгоритми зв'язування даних в причинно-наслідкову модель. Один із напрямків побудови та розвитку цих алгоритмів базується на дослідженні аналогічних біологічних моделей, так як люди легко справляються з такими задачами. Проте дані з нейронауки часто важко формалізувати в обчислювальні моделі, які не тільки будуть відтворювати спостережувані ефекти, але й пояснювати принципи обробки інформації.

Область нейронауки надзвичайно динамічно розвивається через нові технічні та теоретичні засоби дослідження, що виявляється в сотнях тисяч публікацій щороку. Основні напрямки, на які спирається дана робота, та які можуть потенційно породити нові, кращі алгоритми обробки даних включають роботу одного нейрона, математичні моделі якого були покращені такими вченими як E.Izhikevich, W.Gerstner, C.Koh; обчислювальні властивості дендритного дерева, досліджені в роботах B.Mel, P.Poirazi, G.Kastellakis, V.Sjostrom, T. Branco, M. London. В працях B.Olshausen, J.Hawkins, показано важливість розрідженої активації нейронної мережі, що узгоджується з теоретичними моделями асоціативної пам'яті в роботах D.Willshaw, A.Knoblanch.

Моделі асоціативної пам'яті популярні у 80-х роках, актуалізуються знову в роботах A.Graves та G.Wayne, через проблему «катастрофічного забування» в класичних нейронних мережах, що навчаються на основі оптимізації функції похибки. Нещодавно, моделі штучних нейронних мереж з пам'яттю набули особливого інтересу, так як потенційно можуть лягти в основу створення причинно-наслідкової моделі середовища для покращення розпізнавання та прийняття рішень в задачах робототехніки.

Дослідження біологічних нейронних мереж є джерелом багатьох нових ідей і вказують на важливість представлення даних у вигляді розріджених багатомірних векторів для побудови штучних мереж з пам'яттю. Часто розглядається бінарна активація нейронів, зокрема в роботах Д.А.Рачковського, Р. Kanerva. Операції над бінарними векторами, такі як додавання, перетин, асоціювання не тільки швидко виконуються при апаратній реалізації, але й відповідають біологічним експериментальним даним. Тому розробка моделей з таким представленням даних є перспективним напрямком як для кращого розуміння роботи реальних мереж, так і для створення більш ефективних штучних.

Отже, тематика обробки інформації та її збереження в розподіленому вигляді в нейронних мережах швидко розвивається, але існуючі алгоритми все ще далеко позаду природних біологічних систем. Тому, створення та дослідження нових моделей, які враховують останні біологічні дані можуть породити нові, кращі алгоритми та бути застосовані для задач розпізнавання та робототехніки.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана у відділі прикладного нелінійного аналізу Інституту прикладного системного аналізу Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» відповідно до планів наступних науково-дослідних тем:

«Системи, моделі та рівняння з випередженням та запізненням. Дослідження математичних задач та застосувань», (номер державної реєстрації 0114U000205.

«Нові узагальнені математичні моделі з дискретним часом: клітинні автомати, нейромережі, дискретні рівняння та можливі їх застосування» – номер державної реєстрації 0119U002100.

Також робота виконувалася в лабораторії LORIA, м. Нансі, Франція, в рамках програми обміну Erasmus + та коледжі Каламазу, м.Каламазу, США.

Мета і задачі дослідження. Метою дослідження є розробка моделей просторово-часової асоціативної пам'яті в моделях біологічних нейронних мереж; а також, проведення аналізу їх обчислювальних властивостей та можливості їх практичного застосування в задачах робототехніки.

Для досягнення мети дослідження поставлені й розв'язані такі задачі:

1. Побудовано двомірну модель навчання з різними часовими масштабами в спайковій нейронній мережі на основі сучасних даних з нейронауки.
2. Запропоновано нову модель асоціативної пам'яті з використанням моделі нейрона сігма-пай. Отримано аналітичні вирази ємності пам'яті. Порівняно зі стандартною до порівняння моделлю асоціативної пам'яті Вілшова.
3. Застосовано нову модель асоціативної пам'яті для задачі прогнозування послідовностей, проведено теоретичний аналіз ємності пам'яті переходів та порівняно з подібною моделлю НТМ.
4. Побудовано модель представлення послідовності в бінарне розріджено-розподілене представлення на основі біологічних відомостей алокації (allocation) пам'яті.

Об'єктом дослідження є активація та навчання в моделях біологічних нейронних мереж.

Предмет дослідження – моделі асоціативної пам'яті у спайкових та сігма-пай нейронних мережах

Методи дослідження – як теоретичні (системний аналіз, узагальнення) так і практичні (проведення комп'ютерних симуляцій та обчислювальних експериментів). Технічна частина включає застосування теорії ймовірності та комбінаторного аналізу, теорії динамічних систем, теорії інформації та комп'ютерного моделювання.

Наукова новизна отриманих результатів.

Вперше:

- враховано дендритні обчислення для побудови нової моделі асоціативної пам'яті використовуючи модель нейрона сігма-пай, відрізняється тим, що має високу ємність збережених пар векторів;

- показано, що врахування здатності нейрона до активації приводить до часової інтеграції послідовності на масштабі часу секунд, що якісно відтворює біологічні ефекти такі як: збереження схожості, чутливість до порядку, доповнення послідовності та часову схожість.

Удосконалено:

- правила пластичності з різними часовими масштабами, відрізняється тим, що додано вплив повільної змінної сили зв'язку на швидку, що краще відповідає молекулярним механізмам в синапсах;
- ємність пам'яті в моделі прогнозування послідовності з використанням розріджено-розподіленого представлення даних, відрізняється тим, що використовується функція активації сігма-пай та часова затримка в якості контексту.

Набула подальшого розвитку:

- теоретична оцінка ємності пам'яті на основі нейрона з дендритами, відрізняється тим, що досліджується не один нейрон, а мережа нейронів.

Практичне значення одержаних результатів полягає у побудові моделі асоціативної пам'яті високої ємності, яка може бути використана для швидкого пошуку по базі даних, а також у розпізнавання образів в робототехніці. Теоретичний аналіз показує важливість врахування дендритних обчислень в моделі нейрона, що може бути використано практично в побудові нейроморфних архітектур. Часова інтеграція дозволяє представляти послідовність елементів (наприклад, візуальних рис) в окреме представлення, що може бути асоційоване та використане для розпізнавання послідовності. Розроблена схема, що дозволяє навчати робота розпізнавати нові об'єкти в режимі онлайн. Хоча є ще багато невирішених проблем, проте, результати даної роботи мають практичне значення в побудові біологічно подібних алгоритмів навчання.

За результатами стажуванні в лабораторії Loria, м. Нансі, Франція отримано акт впровадження результатів в роботу лабораторії. Також, впроваджено в навчальний процес КПІ ім. Сікорського в курс «Нейронні мережі».

Особистий внесок здобувача. Всі наукові результати, що складають основний зміст роботи та становлять наукову новизну, отримані автором особисто. Більшість публікацій, результати яких увійшли в дану роботу автор опублікував особисто. Виняток становить робота [1] виконані у співробітництві з колегами, в якій автор побудував модель асоціативної пам'яті, провів комп'ютерні симуляції та отримав аналітичні вирази оцінки її ємності.

Апробація результатів дисертації. Наукові та практичні результати доповідались на семінарах та наукових конференціях:

- Осауленко В.М. На шляху до створення інтелекту *Philosophed 2015 міжнародна наукова конференція, ст. 47-49, Київ 2015.*
- Осауленко В.М. Тестування простих моделей з дендритами для розрідженого бінарного представлення зображень. XVII Міжнародна науково-технічна

конференція Штучний інтелект та інтелектуальні системи. Artificial intelligence and intellegent systems (aiis'2017), Київ 2017, с. 101-108

- Osaulenko V.M. New ideas to test for artificial neural networks inspired from neuroscience *Міжнародна наукова конференція «Сучасна інформатика: проблеми, досягнення та перспективи розвитку» Інституту кібернетики імені В.М.Глушкова, 13-15 грудня 2017, с.82-83*
- Osaulenko V.M. Learning of higher-order correlations improves capacity of an association memory. *Nonlinear analysis and applications 4rd International Conference of memory of corresponding member of National Academy of Sciences of Ukraine Valeriy Sergeevich MELNIK (1952 - 2007), NTUU KPI, Kyiv, Ukraine, April 04-06, 2018, p.76*
- Osaulenko V.M. *Model for sequence prediction based on dendritic spatiotemporal integration*. 20-th International conference on System Analysis and Information Technology SAIT 2018, May 21-23, 2018 pp.36-37 Institute for Applied System Analysis of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
- Osaulenko V.M. Dendritic nonlinearity provides high capacity sequence memory *FENS FORUM of neuroscience 2018, Berlin, p.368*
- Makarenko O.S., Osaulenko V.M. Application of cellular automates in some models of artificial intelligence. Конференція IEEE SAIC-2018, 2018, 1-4. 10.1109/SAIC.2018.8516837
- Науковому семінарі при Навчально-науковому комплексі "Інститут прикладного системного аналізу" КПІ ім. Ігоря Сікорського (Київ, 2018).

Публікації. Основні результати дисертаційної роботи опубліковано в 13 наукових працях у тому числі 6 статей у наукових фахових виданнях (з них 2 статті у виданнях іноземних держав, 4 статті у виданнях України, які включені до міжнародних наукометричних баз), 7 тез доповідей в збірниках матеріалів конференцій.

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох основних розділів, висновків, списку використаних джерел та додатків. Робота викладена на 158 сторінках і містить 132 сторінок основної частини, 50 рисунків, 1 таблицю і список використаних джерел із 177 посилань.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність розробки біологічно подібних алгоритмів просторово-часової асоціативної пам'яті, визначені мета та задачі, об'єкт, предмет і методи дослідження; наведено зв'язок з науковими програмами; зазначено наукову новизну та практичне значення одержаних результатів, показано особистий внесок здобувача.

У **першому розділі** наведено огляд літератури за тематикою обчислювальної нейронауки та штучних нейронних мереж. Дано короткий опис роботи біологічного нейрона, його динаміки та правил пластичності. Описано основні

характеристики біологічних нейронних мереж та шляхи поширення інформації між областями мозку. Підкреслено важливість проблеми кодування інформації в біологічних нейронах та способів представлення даних, таких як розріджені та щільні бінарні вектори. Показано основні проблеми навчання в моделях біологічних нейронних мереж. Дано пояснення просторово-часової асоціативної пам'яті, яка в загальному визначається як відображення однієї послідовності векторів в іншу. Випадок, коли послідовність складає всього лише один вектор, називається просторовою пам'яттю (вектор як одновимірний простір). Часова пам'ять розглядає дані розподілені в часі, які представляються послідовністю векторів. Часто дані мають і просторову і часову складові, наприклад, звук, тому виникають саме просторово-часові асоціації.

В другій частині даного розділу беруться до уваги останні результати дослідження взаємодії молекулярних механізмів пластичності в синапсах між нейронами. На основі цього побудовано двовірну модель навчання в спайковій нейронній мережі, що дозволяє поєднати різні правила пластичності. Відрізняється від інших тим, що в явному виді вводиться окрім звичайної сили зв'язку w , додаткова змінна w_s , яка враховує зміну сили зв'язку на повільнішому масштабі часу. Ці дві змінні взаємодіють між собою через систему зв'язаних диференціальних рівнянь, що дозволяє формувати стабільні асоціативні зв'язки:

$$\begin{cases} \frac{dw^i}{dt} = \frac{1}{\tau_w} (w_s^i - w^i) + \frac{1}{\tau_w} (S^i(t)(x^i \cdot p_{slow}) - S_{post}(t)p) \\ \frac{dw_s^i}{dt} = \frac{1}{\tau_{w_s}} (w^i - w_s^i)M + \frac{1}{\tau_{w_s}} w_s^i (1 - \frac{y^s}{A})^3 \end{cases} \quad (1)$$

Де w^i, w_s^i швидка та повільна змінна відповідно, τ_w, τ_{w_s} часові сталі, що визначають швидкість зміни ваг. Модель також враховує стандартне хебівське правило триплетної спайк-часово-залежної пластичності (другий доданок першого рівняння) та гомеостатичну пластичність (другий доданок другого рівняння). Додаткові змінні x^i, p, p_{slow} – низькочастотні фільтри активації нейрона, y^s середня частота активації нейрона, A – константа, M – сигнал підкріплення. $S^i(t) = \sum_f \delta(t - t_f^i)$ де δ -та – функція Дірака та t_f^i – час надходження спайку.

Поведінку моделі схематично зображено на рис. 1 за умови відсутності хебівського та гомеостатичного доданків справа. Розглянемо один момент навчання, коли відповідно до хебівської пластичності w зросла (Рис.1А). Тоді, згідно рівнянь (1), w починає зменшуватись, а w_s зростає. Загалом, повільна змінна збільшиться на Δw , що залежить від співвідношення τ_w / τ_{w_s} . Якщо навчальна подія повториться, тоді загальний ріст Δw буде більшим (Рис.1В). Це добре відповідає твердженню, що реактивація навчання покращує пам'ять. У третьому випадку, нейрон отримує сигнал підкріплення $M \gg 1$ в момент часу відмічений як *. Тоді швидкість зміни w_s буде значно більша і $w_s \rightarrow w$, що виразиться у великій зміні Δw (Рис.1С).

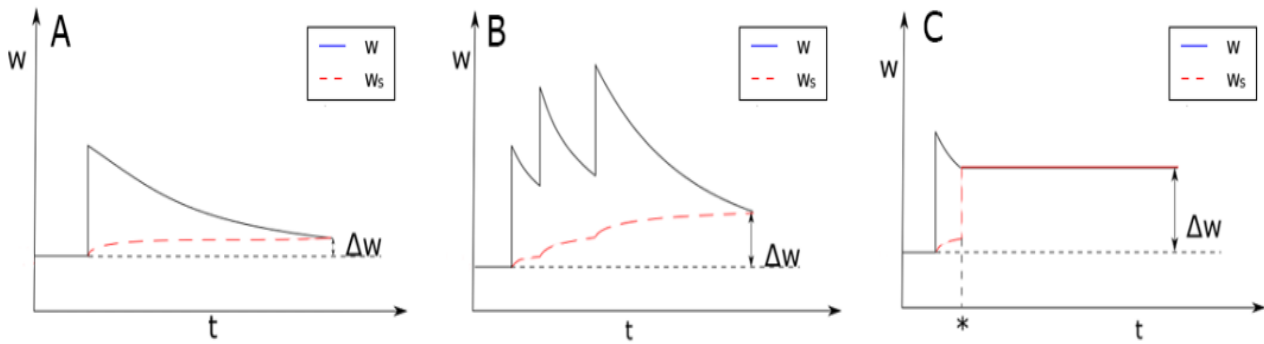


Рисунок 1. Зображення еволюції швидкої та повільної змінних. А: зображення динаміки двох змінних при одному моменті навчання. **В:** повторення навчання, призводить до збільшення сили зв'язку. **С:** сигнал підкріплення M , майже миттєво збільшує силу зв'язку

Для тестування моделі в мові програмування Python використовувався один вихідний нейрон, який отримував сигнал від 100 вхідних нейронів. Ваги приймали тільки додатні значення, тому вхідні нейрони тільки збуджували вихідний. В якості моделі нейрона використано стандартну дисипативну модель інтегруй-активуй, що описує зміну напруги на мембрані. Для перевірки навчання виділялися групи (паттерн) з 10-15 нейронів з більшою частотою активації і аналізувалася зміна сили зв'язку.

Один із основних результатів представлено на рис. 2. На графіках зліва, зверху до низу показано усереднено активацію вихідного нейрона, усереднену активацію на більшому масштабі часу, та динаміку зміни w, w_s для одного зв'язку. На рисунку справа, показано зміну w_s для всіх 100 зв'язків і стрілками позначено збільшені сили зв'язку внаслідок навчання, що відповідають частішому паттерну активації. Також, збережено інші паттерни, хоча і з меншим значенням сили зв'язку.

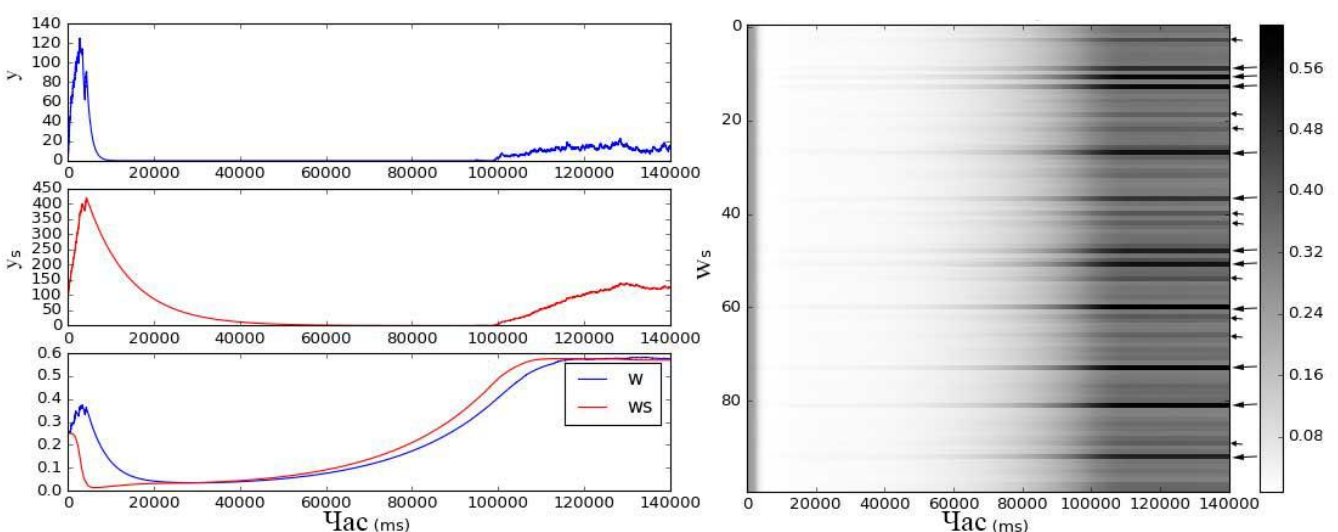


Рисунок 2. Поведінка моделі при активації кількох паттернів. Великі стрілки вказують на частіший паттерн. Малі стрілки вказують на інші менш частіші паттерни, які зберегли своє представлення.

Подібно до попередніх, дана модель дозволяє виділяти більш статистично вагоміші стимули, тим самим реалізуючи простий варіант навчання без учителя. Також зберігається стабільність цих зв'язків при спонтанній активації мережі.

Запропонована модель має ряд відмінностей від попередніх підходів.

1) У явному вигляді вводиться повільна змінна w_s . Припускається, що w_s відповідає розміру шипика і молекулярній машинерії, що підтримує його стан, включаючи регуляторні протеїни, такі як СаМКІІ та РКА. Швидка змінна w відповідає кількості та ефективності АМПА рецепторів.

2) Модель природньо включає навчання з підкріпленням через змінну M . Збільшення M , змінює відношення τ_w / τ_{ws} , що змінює швидкість навчання. Це може відповідати впливу нейромодуляторів на швидкість молекулярних механізмів.

Таким чином, модель зберігає просторові паттерни активності зі збереженням стабільності зв'язків та мережі в цілому. Тим не менш, робиться висновок, що практичне застосування, зокрема в задачах розпізнавання, даної моделі і взагалі моделей обчислювальної нейронауки не можлива на даному етапі. Досі стоїть проблема необхідного масштабу опису, наприклад, чи потрібно враховувати допорогові флуктуації напруги, чи електричні властивості дендритів, чи, можливо, потрібно робити опис на молекулярному рівні, виділяючи концентрації різних видів молекул, кожна зі своєю роллю, чи навіть варто все моделювати на генетичному рівні? Тому, наступні розділи використовують більш спрощені абстрактні моделі для дослідження обчислювальних властивостей асоціативної пам'яті, при цьому зменшуючи біологічну деталізацію.

Другий розділ присвячений просторовій асоціативній пам'яті, що визначається як відображення f між R бінарними векторами $\{f : \mathbf{x}^\mu \rightarrow \mathbf{y}^\mu, \mu = 1 : R\}$ де R – кількість асоційованих пар, а $\mathbf{x}^\mu, \mathbf{y}^\mu \in \{0,1\}^N$ – бінарні вектори розмірності N . У випадку коли вхідний і вихідний вектори різні $\mathbf{x} \neq \mathbf{y}$ пам'ять називається гетеро-асоціативною, і класична модель в цьому випадку – модель Вілшоу. Якщо вектори однакові $\mathbf{x} = \mathbf{y}$, то пам'ять авто-асоціативна, і класична модель – модель Хопфілда.

Більшість асоціативних моделей під впливом моделі нейрона Маккалоха і Піттса та правила навчання Хебба використовують зважену сумачію вхідного вектору $f(\sum_i w_{ik} x_i)$, де параметри w_i навчаються через відслідковування парних кореляцій $\Delta w \sim x_i \cdot y_k$. В даній роботі наводиться основний недолік такого підходу – часте виникнення помилкових асоціацій. Раніше був запропонований можливий спосіб вирішення цього недоліку через відслідковування збігів активації вищих порядків $x_i x_j \dots y_k$. Проте, він обчислювально складний, так як включає багатовимірні матриці для повнозв'язної мережі.

На основі даних дослідження реальних біологічних нейронів в даній роботі показано, що зменшення обчислювальної складності зі збереженням низького рівня помилкових асоціацій, може бути досягнуто завдяки врахування

дендритних обчислень в моделі нейрона. Раніше було показано, що дендритні сегменти інтегрують інформацію нелінійно, і якщо в короткий проміжок часу відбувається локальна активація кількох синапсів, то сумарна активація більша, ніж сума окремих активацій. Біологічний механізм такого ефекту включає нелінійні властивості NMDA каналів та Ca^{2+} залежних каналів. У такому випадку може виникнути дендритний спайк, який менше затухає і відповідно активує нейрон сильніше. Обчислювально, така нелінійна сумація відповідає детектору збігів, коли визначається селективна активація певної комбінації нейронів. Враховуючи, що середня довжина дендритного дерева типового пірамідального нейрона в корі головного мозку порядку 10^4 мкм, а дендритний спайк може виникнути на сегменті розміром порядку 10 мкм, можна говорити, про множинне детектування збігів одним нейроном.

Раніше було враховано дендритні обчислення в модель нейрона сігма-пай (2), і в даній роботі вперше було побудовано та досліджено асоціативну пам'ять з таким нейроном. Для цього було використано спрощений алгоритм навчання (3).

$$y_j = \theta\left(\sum_m \prod_{i \in G_m^j} x_i\right) \quad (2)$$

$$\forall j: y_j = 1, \begin{cases} G_{m_j+1}^j = \{i \mid i \in A_x^k\} \\ m_j = m_j + 1 \end{cases} \quad (3)$$

Де $\theta(x) = \{1, x > 0, \text{інакше } 0\}$ - крокова функція, $G_{m_j}^j$ - m_j дендритний сегмент або кластер j -го нейрону. При навчанні для кожного активного нейрону у векторі \mathbf{y} вибираються k індексів активних вхідних нейронів і записується в новостворений кластер. Таким чином якщо всі k вхідних нейронів, що записані в кластер активні $x_i = 1$, тоді добуток ненульовий і вихідний нейрон активується $y_j = 1$. Схематично модель нейрона та навчання зображена на рис.3.

Було проведено теоретичний аналіз ємності пам'яті та виведено формулу максимальної кількості збережених асоціацій у граничному випадку $1/N \rightarrow 0$:

$$R_{\max} = \frac{\varepsilon}{s(1-s)s'} \quad (4)$$

Де $s = \frac{a}{N}$ - розрідженість активації вектору \mathbf{y} , та $s' = \frac{a'}{N} = \frac{C_a^k}{C_N^k}$, де a - кількість

одиниць, або активних нейронів з поміж всієї кількості N , $a' = C_a^k$, $N' = C_N^k$ -

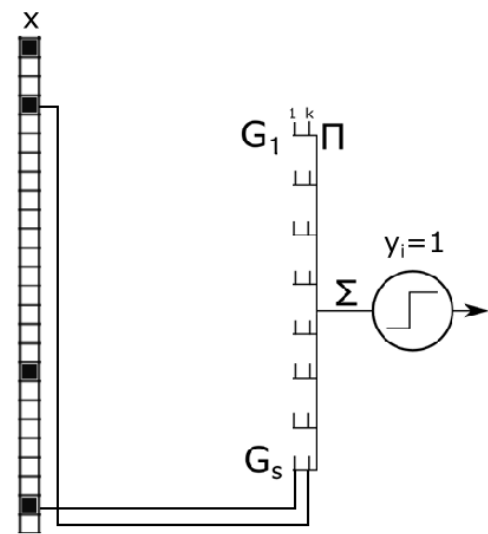


Рисунок 3. Модель нейрона з нелінійними дендритами та схема навчання, як створення зв'язків до активних нейронів.

комбінаторне розширення вхідного вектору за рахунок детектування збігів активації. Також задається точність асоціативної пам'яті, наприклад, $\varepsilon = 0.01$, що означає, що не допускається більше, ніж 1% помилок в отриманому вихідному векторі \mathbf{y} . Цікаво, що теоретичний результат показав, що модель нейрона з дендритами комбінаторно збільшує розмірність вхідного вектору, в якій легше розділити асоціації.

Для підтвердження аналітичного виразу максимальної кількості збережених асоціацій були проведені комп'ютерні симуляції, де згенеровано пари векторів $(\mathbf{x}^\mu, \mathbf{y}^\mu)$, $\mu = 1:R$. У фазі навчання застосовується рівняння (3), у фазі активації застосовується рівняння (2), де отримується вектор $\tilde{\mathbf{y}}$ який порівнюється із збереженим \mathbf{y} . Результати показують узгодження теоретичного та експериментального значення максимальної кількості збережених асоціацій.

Також, було досліджено залежність максимальної кількості збережених асоціацій в залежності від параметрів моделі. Показано, що максимальна ємність досягається за розрідженої активації коли лише мала частина нейронів активна відносно загальної кількості. Результати представлено на рис.4

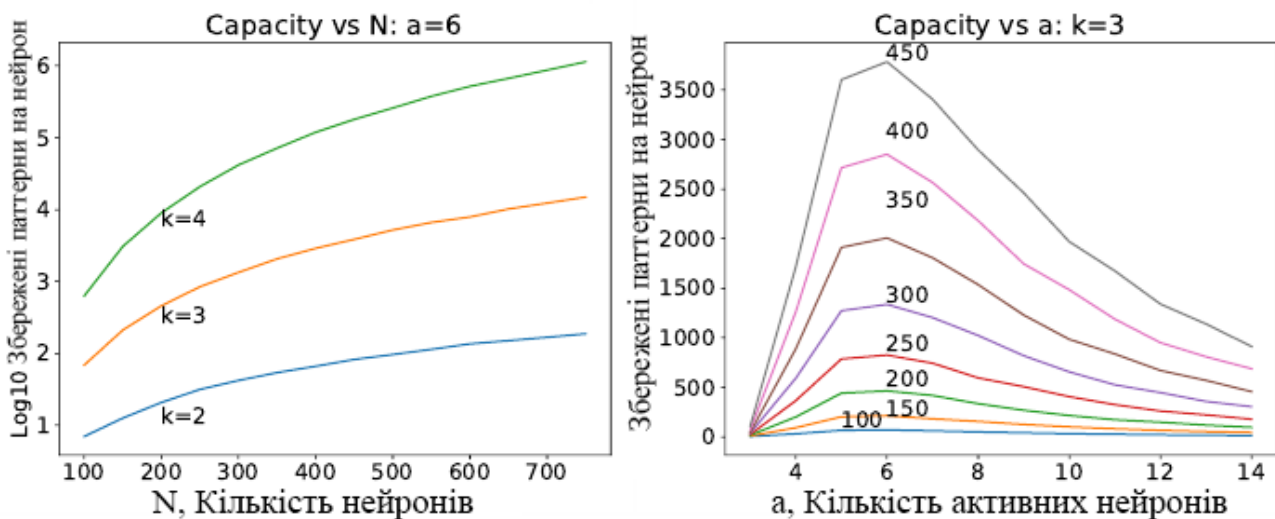


Рисунок 4. (Зліва) Залежність десяткового логарифму кількості збережених асоціацій на нейрон відносно кількості нейронів (розмірності вектору). Один нейрон може зберігати тисячі окремих паттернів активації. (Справа) Залежність збережених асоціацій на нейрон відносно кількості активних нейронів. Як видно, є оптимальне значення за якого ємність максимальна.

Максимальна кількість збережених асоціацій не може бути використана для порівняння різних моделей пам'яті, так як потрібно ще враховувати кількість параметрів моделі (зв'язків між нейронами). Для цього було застосовано стандартні міри ємності збереженої інформації: мережеву, синаптичну та інформаційну ємності. В таблиці 1, наведені означення та назви мір, а також формули обрахунку для запропонованої моделі. Мережева ємність визначається як відношення ємності каналу до кількості зв'язків. $T(N, a, R_{\max}) = \sup I(\tilde{\mathbf{y}}, \mathbf{y})$ -

ємність каналу, яка обраховується як максимальна взаємна інформація між відтвореним та збереженим векторами. Синаптична ємність враховує кількість

ненульових зв'язків, а інформаційна ємність враховує кількість інформації на один зв'язок. Для умови $a \ll N$ можна обрахувати наближення $T(N, a, R_{\max}) = R_{\max} N H_2(\frac{a}{N}) \approx R_{\max} a \log_2(\frac{a}{N})$. Також, враховуючи, що в моделі нейрона з дендритами за R асоціацій утворюється Rka зв'язків, звідси і отримуємо значення різних мір ємності для запропонованої моделі.

Назва міри	Формула означення	Ємність моделі
Мережева ємність	$C^N = \frac{T(N, a, R_{\max})}{\# \text{ зв'язків}} [\text{bit} / \text{contact}]$	$C^N = \frac{1}{k} \log_2(\frac{N}{a})$
Синаптична ємність	$C^S = \frac{T(N, a, R_{\max})}{\# \text{ ненульових зв'язків}} [\text{bit} / \text{synapse}]$	$C^S = \frac{1}{k} \log_2(\frac{N}{a})$
Інформаційна ємність	$C^I = \frac{T(N, a, R_{\max})}{\# \text{ bits per contact}}$	$C^S = \frac{\log_2(\frac{N}{a})}{k \log_2(N)}$

Таблиця 1. Міри ємності та формули для обрахунку для запропонованої моделі асоціативної пам'яті.

На рис.5 показано порівняння синаптичної ємності для запропонованої моделі та для стандартної до порівняння мережі Вілшоу. Як видно з графіку, за параметрів запропонованої моделі $k=2$ та $k=3$ є перевага в синаптичній ємності для $R=380$ в 39.8% та 71.4% відповідно. Варто зазначити, що дана модель має нижчу інформаційну ємність, що пов'язано з необхідністю зберігати індекс синапса на дендритному дереві в, щонайменше, $\log_2(N)$ -бітовому числі. Проте, для практичної реалізації пам'яті у нейроморфних архітектурах, важливою є саме синаптична ємність, яка враховує кількість необхідних фізичних зв'язків.

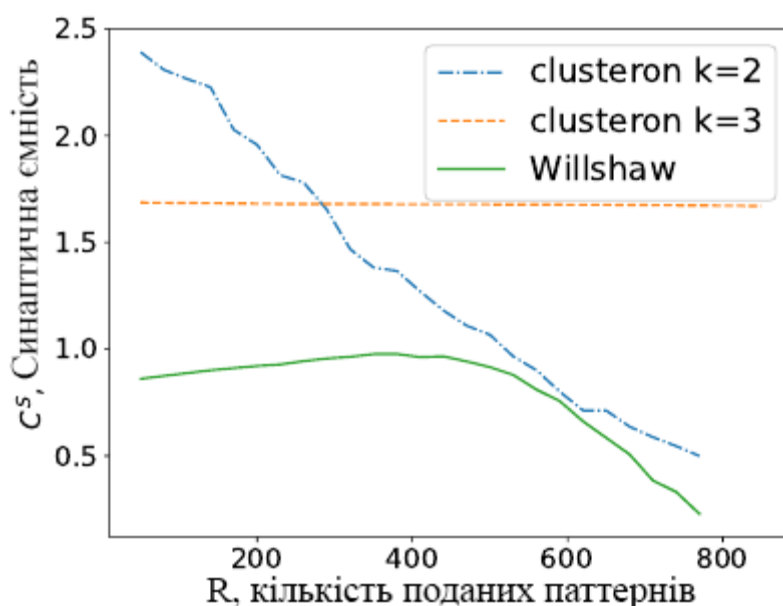


Рисунок 5. Порівняння синаптичної ємності збереженої інформації запропонованої моделі для різних параметрів та моделі Вілшоу.

Загалом, модель просторової асоціативної пам'яті відповідає біологічним уявленням про роботу дендритного дерева як детектора збігів. І було показано, що такий новий біологічний підхід має обчислювальні переваги відносно лінійної сумачії входної активації.

В **третьому розділі** розглянуто часову асоціативну пам'ять, що в загальному з'єднує послідовність входних векторів з послідовністю вихідних.

Спершу, досліджується задача прогнозування послідовності, що може бути формалізована як відображення $f: \mathbf{x}_0, \dots, \mathbf{x}_{j-1} \rightarrow \mathbf{x}_j$, що видобуває наступний елемент послідовності, де $\mathbf{x}_i \in \{0,1\}^N$ - елемент послідовності представлений як бінарний вектор розмірності N . Використовується відомий підхід часової затримки, коли останні T елементи послідовності перетворюються в один великий вектор шляхом конкатенації (рис 6А.). Тоді, задача зводиться з часової до просторової асоціативної пам'яті представленої в другому розділі. Проте, для успішного прогнозування потрібно обов'язково включати активні нейрони в кластер з останнього елементу послідовності, інакше буде видобуватися не наступний, а більш майбутній елемент.

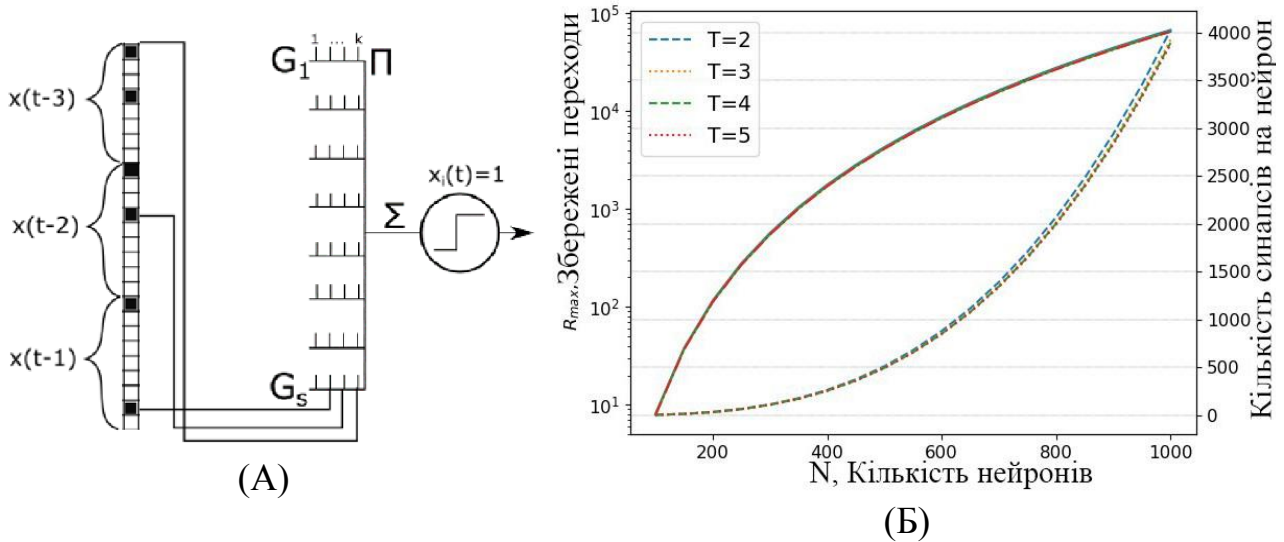


Рисунок 6. (А) Ілюстрація ідеї створення зв'язків з минулими активними нейронами. (Б) Максимальна кількість переходів послідовностей для різних розмірів мережі та кроків часу показана як верхня лінія. Нижня лінія показує кількість зв'язків на один нейрон

Проведено теоретичний аналіз ємності кількості переходів даної моделі подібно до просторової асоціативної пам'яті. Якщо прийняти, що створюється k зв'язків до передостаннього елементу, та m зв'язків до елементів з проміжку $[t-T, t-2]$, то максимальна кількість збережених переходів визначається рівнянням (4), з модифікацією $N' = C_N^k C_{N(T-1)}^m$ та $a' = C_a^k C_{a(T-1)}^m$. На рис.6Б. показані результати теоретичного аналізу моделі, де показано залежність максимальної пам'яті переходів від розміру мережі N для різних часових вікон. Як видно, ємність не залежить від часового вікна T і збільшується з розміром мережі. Для $N=1000$ потрібно 4000 синапсів на нейрон, що відповідає біологічним даним, а кількість переходів дорівнює 10^5 . Якщо взяти послідовності букв, що утворюють слова з

середньою кількістю символів 7, мережа може записати та спрогнозувати близько 14000 можливих слів. Варто зазначити, що вибрано розріджено-розподілене представлення елементів послідовності, так як це забезпечує максимальну ємність пам'яті.

Було проведено порівняння даної моделі з відомою моделлю ієрархічно темпоральної пам'яті Хоккінса (HTM). Показано, що модель зберігає всі попередні переваги, зокрема, навчання онлайн, використання SDR (розріджено-розподілене представлення), формування багатьох одночасних прогнозів, використання локальних правил навчання та стійкість до шуму. Основна відмінність у способі збереження контексту. В моделі HTM, контекст послідовності визначається через активність нейронів в колонках (спеціальній архітектурі мережі) і зберігається протягом всієї послідовності. В запропонованій моделі, контекст визначається завдяки затримці активності в часовому вікні T . Перевага такого підходу в спрощенні архітектури мережі, недоліком є неможливість робити передбачення далі часового вікна T . Проте, поєднуючи дану модель з моделлю часової інтеграції представленою далі, цей недолік можна усунути. Також, порівнюючи з HTM, використання мультиплікативної активаційної функції дає перевагу у більшій ємності пам'яті при однаковій кількості зв'язків. Так, для попереднього прикладу $N=1000$ та 4000 синапсів на нейрон, ємність моделі HTM складає $\sim 10^4$ переходів.

Загалом, запропонована модель являє собою базовий і простий спосіб прогнозування елементів послідовності з використанням розріджено-розподілених представлень та локального правила навчання. Головна мета моделі виконана – запропонувати можливий спосіб прогнозування послідовностей в біологічних нейронних мережах з урахуванням сучасного розуміння дендритних обчислень.

Друга частина третього розділу доповнює модель часової асоціативної пам'яті шляхом розв'язку задачі часової інтеграції, що може бути представлена як пошук відображення $f: \{\mathbf{x}_t\} \rightarrow \mathbf{y}$, де $\mathbf{x}_t \in \{0,1\}^N$ - елемент послідовності представлений як бінарний вектор розмірності N , $\{\mathbf{x}_t\} = \mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_t \in \{0,1\}^{t \times N}$ послідовність елементів, $\mathbf{y} \in \{0,1\}^M$ -вектор, що інтегрує всю послідовність, який має ряд властивостей показаних далі. Задача прогнозування та задача часової інтеграції разом вирішують задачу навчання послідовностей, яка є фундаментальним обчислюванням, що виконується в біологічних нейронних мережах.

В даній роботі запропонована нова біологічно подібна модель, яка виконує часову інтеграцію і представляє послідовність у бінарний розріджено-розподілений вектор. В основу моделі покладена ідея алокації пам'яті в нейрони зі більшою здатністю до активації в поєднанні з розрідженим представленням даних. Суть ідеї полягає в тому, що нейрони мають різний рівень здатності до активації, тобто готовності генерувати потенціал дії, в залежності від попередньої динаміки мережі. Тому, вхідні зв'язки на нейронну мережу сформуються з нейронами зі збільшеною здатністю до активації, а отже до навчання. В такі нейрони і записується (алокується, від англ. allocation) пам'ять.

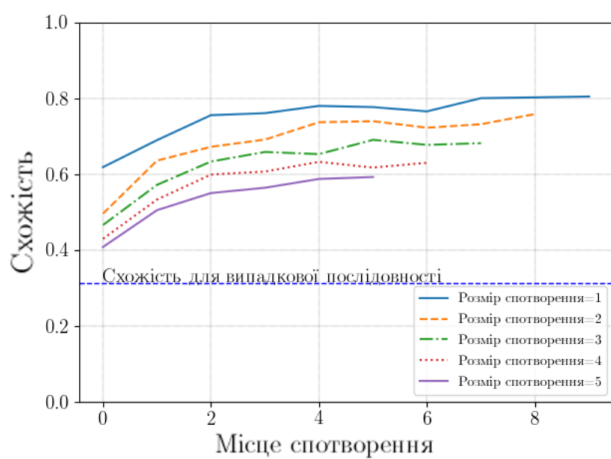
Для вирішення задачі побудована нейронна мережа з трьома шарами прямого поширення. На вхід подається послідовність $\alpha_1\alpha_2..\alpha_T$ довжиною T по одному елементу. Елементи послідовності належать до скінченного словника $\alpha_i \in A, |A| < \infty$ (наприклад, букви). Кожному елементу співвідноситься бінарний код $\alpha_i \rightarrow \mathbf{x}_i \in \{0,1\}^N$. Першим шаром мережі слугує вектор \mathbf{x} . Другий шар $\mathbf{z}_i \in \{0,1\}^M$ переводить вхідний вектор в бінарний розріджений вектор через:

$$\begin{cases} \mathbf{z}_i = k\text{WTA}((1 + \mathbf{v}_i)(\mathbf{W} \cdot \mathbf{x}_i), k_1) \\ \mathbf{v}_{i+1} = \mathbf{v}_i + \tau(\mathbf{z}_i - \mathbf{v}_i) \end{cases} \quad \begin{matrix} (5) \\ (6) \end{matrix}$$

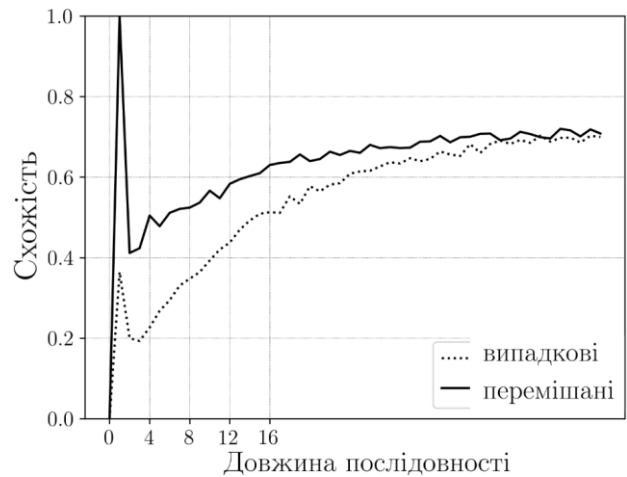
Де $\mathbf{W} \in \{0,1\}^{M \times N}$, $w_{ij} \in \text{Ber}(s_w)$ матриця випадкових бінарних ваг з розподілу Бернуллі при $M \gg N$, $k\text{WTA}(\cdot)$ - функція (k-winners-take-all), яка вибирає k_1 найбільших компонентів вектору і повертає їх значення одиницями, а всі інші є нульові. Така процедура подібна до алгоритму випадкової проекції, який формує розріджені бінарні вектори, де зберігається схожість між вхідними векторами $\mathbf{x}_i, \mathbf{x}_j$ та їх проекціями $\mathbf{z}_i, \mathbf{z}_j$. В даній роботі ця процедура видозмінена шляхом множення лінійної сумачі в рівнянні (5) на здатність нейронів до активації \mathbf{v}_i , що визначається як низькочастотний фільтр (або рухоме середнє) від \mathbf{z}_i через рівняння (6) з параметром τ .

Після того як кожний елемент послідовності спроектований на свій вектор \mathbf{z}_i , визначається вектор $\mathbf{y} \in \{0,1\}^M$, що представляє всю послідовність:

$$\mathbf{y} = k\text{WTA}\left(\sum_{i=1:T} \mathbf{z}_i, k_2\right) \quad (7)$$



(А)



(Б)

Рисунок 7. (А) Графік залежності схожості послідовності з викривленою копією в залежності від місця викривлення. Різні криві показують різний ступінь викривлення. (Б) Графік залежності схожості послідовності з послідовністю з переставленими елементами відносно довжини послідовності. На нижній кривій зображено схожість з випадковою послідовністю.

Застосовується також функція $kWTA(.)$ для суми векторів \mathbf{z}_i для всіх елементів послідовності. Тільки в цьому випадку вибирається менша кількість активних нейронів, що закодують всю послідовність $k_2 < k_1$.

Особливість представлення вектору \mathbf{u} в тому, що воно відтворює біологічні ефекти такі як: збереження схожості, чутливість до порядку, доповнення послідовності та часову схожість. Для дослідження властивостей вектору \mathbf{u} було проведено ряд експериментів, які порівнювали представлення однієї послідовності \mathbf{u} з представленням іншої $\hat{\mathbf{u}}$ через подібність косинусів. На рис.7(А) показано порівняння представлення \mathbf{u} оригінальної послідовності $\{\mathbf{x}_t\}$ та представлення $\hat{\mathbf{u}}$ спотвореної $\{\hat{\mathbf{x}}_t\}$. Різні криві показують різний розмір спотворення. Як видно, чим більше спотворення, тим менша схожість. Чим ближче спотворення до початку послідовності, тим менша схожість, що пояснюється зміною контексту для наступних елементів. Ключовим є те, що всі спотворенні послідовності мають більшу схожість, ніж довільна випадкова послідовність, яка показана штриховою горизонтальною лінією. Це свідчить про стійкість представлення послідовності до шуму.

На рис.7Б показано порівняння представлення послідовностей з перестановками елементів. Суцільною лінією показано схожість представлення послідовності з представленням цієї ж послідовності, в якій випадковим чином перемішані елементи. Штрихом показано схожість до послідовності з випадковими елементами зі словника. Для коротких послідовностей видно, що послідовність з перемішаними елементами має більшу схожість, ніж випадкова послідовність. Тобто, зберігається схожість при перестановках, але, все одно, послідовності ABC та ACB представлені по різному.

Інший ефект, який спостерігається в даній моделі – це часова схожість. Для послідовності довжиною $T=15$ порівнювалося представлення окремих елементів \mathbf{z}_i послідовності. На рис.8 показана матриця схожості i -го елементу з j -тим. Діагональні елементи показують схожість самих з собою і тому їх значення є одиницями. З рисунка видно, що події, які складають одну послідовність кодуються схожим чином, і чим ближче елементи в часі, тим більш вони схожі. Також, елементи далі в послідовності кодуються більш схоже, ніж на початку. Це пов'язано з тим, що з довжиною послідовності росте здатність до

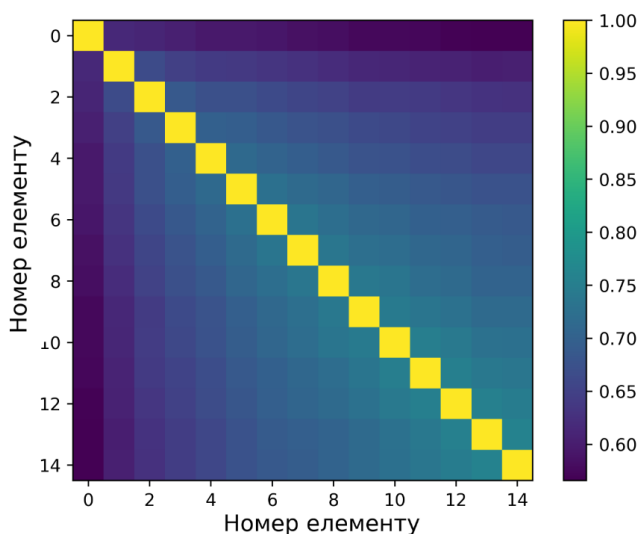


Рисунок 8. Перетин популяцій, що кодують індивідуальні події для послідовності з 15 елементами

активації певної популяції клітин, яка з більшою ймовірністю починає кодувати всі наступні елементи

Загалом, проведені експерименти показали, що дана модель зберігає схожість кодувань для схожих послідовностей, має чутливість до порядку подій, має високу схожість для підпослідовностей та проявляє часову схожість представлення близьких в часі елементів послідовності. Всі ці ефекти спостерігаються в біологічних нейронних мережах, тому є підстави припускати про важливість алокації пам'яті в нейрони зі збільшеною здатністю до активації для задачі часової інтеграції.

Важливо, що вектор u , що представляє всю послідовність може бути далі асоційованим з іншими векторами, за алгоритмом просторової асоціативної пам'яті. Таким чином реалізується асоціювання послідовності до послідовності, що може бути використано для розпізнавання. Іншою особливістю є те, що вектор u може бути використаний в якості контексту для попередньо розглянутої задачі прогнозування. Теоретично, це дозволяє вийти за межі часового вікна T . Більше того, використавши запропонований алгоритм інтеграції подаючи на вхід вектори u можна створити ієрархічні архітектуру для контексту на більшому масштабі часу. Для реалізації такого підходу потрібні деякі модифікації до представлених моделей, які можуть бути додатково досліджені та перевірені в подальшому. Основна ціль даної частини досягнуто – побудовано біологічно подібну модель представлення послідовності на основі біологічних відомостей алокації пам'яті, яка якісно повторює біологічні ефекти.

В четвертому розділі показано зв'язок вирішених задач на прикладі системи розпізнавання просторово-часових паттернів в роботі Nao.

Гуманоїдний робот Nao зображений на рис.7 має висоту 58см та 25 ступенів волі, що дозволяє йому ходити, обстежувати сцену рухами голови, виконувати обмежені дії руками. В стандартній операційній системі Naoqi закладено можливість розпізнавати образи через візуальний вхід з двох камер розташованих на голові. Образи до розпізнавання були попередньо навчені і їх кількість обмежена. Передбачена процедура, як користувач може додати новий образ до розпізнавання. Новий об'єкт поміщається в поле зору робота, об'єкт фотографується, далі користувач повинен вручну виділити контури об'єкта і дати йому мітку(через інтерфейс користувача на комп'ютері до якого попередньо підключений робот). Для кращого розпізнавання, бажано повторити процедуру для різних сторін об'єкта. Очевидно, що така процедура занадто повільна та трудомістка і потребує вдосконалення.



Рисунок 9.
Робот NAO

Сучасні найбільш ефективні системи розпізнавання зображень на основі згорткових нейронних мереж не можуть, поки що, бути використані для навчання в роботах, так як вони потребують обчислювально затратної процедури перенавчання при додаванні нового класу та великої кількості додаткових зображень цього класу.

Моделі просторово-часової асоціативної пам'яті реалізують алгоритми навчання наближені до біологічних, які можуть навчатися онлайн з малої кількості прикладів. Просторова асоціативна пам'ять належить до задачі розпізнавання статичних образів, наприклад, зображень. Часова асоціативна пам'ять належить до задачі розпізнавання рухомих образів, тобто часових послідовностей.

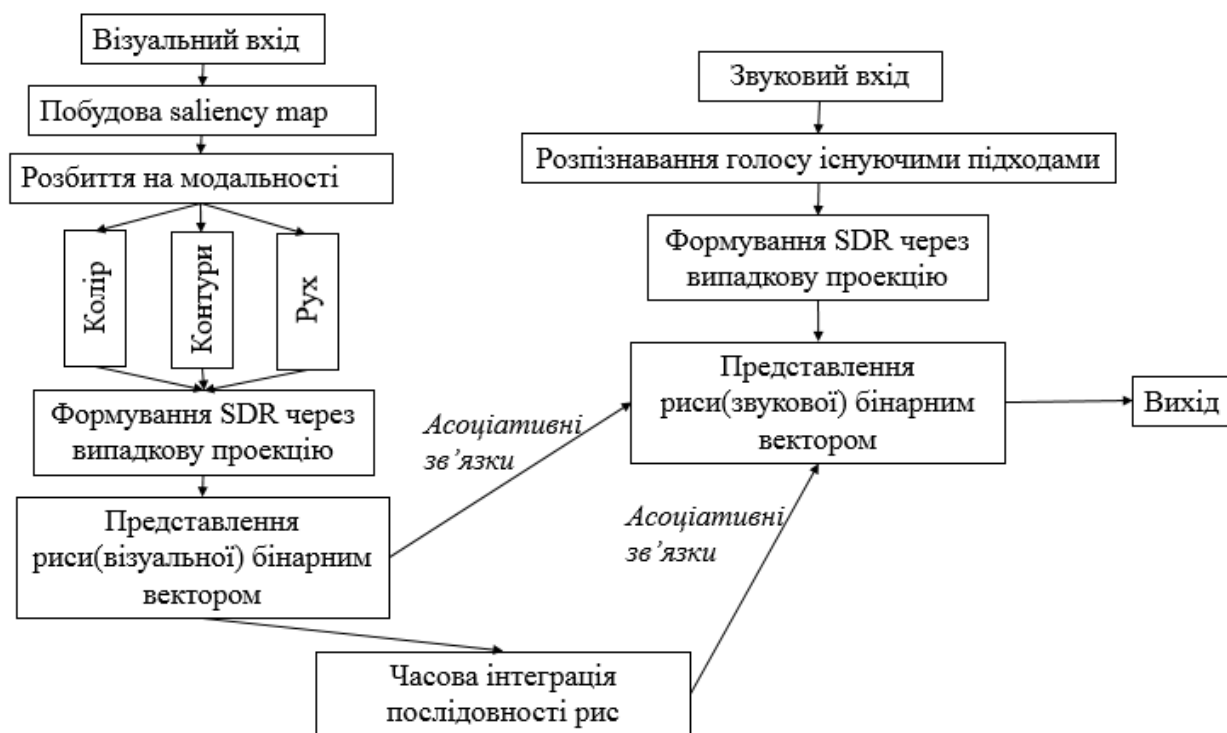


Рисунок 10. Запропонована схема навчання онлайн робота NAO

В роботі запропонована нова схема на рис.10 навчання робота онлайн, що використовує представлені моделі. Вона включає два сенсорних канали, візуальний через камеру та звуковий через мікрофон. Спочатку формується карта виразності (saliency map), коли вибираються найбільш інформативні риси з візуального потоку. Відбувається виділення кольору, контурів, руху подібно до відомих підходів побудови карт виразності. Визначається найбільш виразне місце на зображенні і туди робот повертає голову і центрує погляд. Риса в центрі зору (наприклад, колір об'єкту) через алгоритм випадкової проекції перетворюється у бінарний вектор x . Звуковий сигнал подається на вхід до системи розпізнавання голосу (наприклад, на основі LSTM), що видає текст, який далі також через механізм випадкової проекції перетворюється у бінарний вектор y . Два вектори асоціюються за запропонованим алгоритмом в другому розділі. Загалом, навчання виглядає наступним чином: користувач показує роботу новий об'єкт, рухає ним привертаючи увагу. Робот центрує погляд на об'єкт та перетворює візуальну інформацію в бінарний вектор. Одночасно користувач говорить назву об'єкта, яка розпізнається і також перетворюється в бінарний вектор. Між двома векторами створюються асоціативні зв'язки. Якщо потрібно розпізнати не тільки об'єкт, але й рух, наприклад, помах руки вліво, то використовується алгоритм часової інтеграції, запропонований в третьому

розділі. Послідовність рис, представлених у вигляді бінарних векторів інтегрується в новий бінарний вектор, який асоціюється із звуковою міткою. В майбутньому, коли робот побачить знову цей об'єкт або рух, через асоціативні зв'язки видобувається відповідна мітка **y**, що і буде виходом алгоритму. В загальному випадку вихід використовується для поведінки робота, наприклад, синтез тексту назви об'єкту.

Загалом, представлений підхід дозволяє навчати робота онлайн. Врахування дендритів в модель нейрона забезпечує високу ємність пам'яті. Часова інтеграція перетворює послідовність рис в один вектор, який також може бути асоційованим. Дана схема, в дещо спрощеному вигляді, була реалізована на роботі NAO при стажуванні в лабораторії LORIA у Франції, і показано її життєздатність (відео демонстрація за посиланням www.bit.do/robot_nao) та переваги відносно стандартного способу навчання робота передбаченого його операційною системою. Проте, таке навчання використовує рукотворні риси (handcrafted features) і точність розпізнавання набагато гірша порівняно з методами DeepLearning. Наприклад, на датасеті iCubWorld Transformations Dataset глибокі нейронні мережі мають 70-90% точності (в залежності від мережі та кількості класів), хоча й в реальному житті таке навчання важко назвати онлайн. В той час, точність біологічно подібних алгоритмів близько 10-50%, і сильно залежить від методів кодування. Існує ще багато невирішених задач, зокрема алгоритм випадкової проекції має бути замінений на більш ефективніші методи нейронного кодування. Загалом, є багато майбутніх перспективних напрямків і дана робота є лише маленьким внеском до глобальної задачі побудови розумних роботів.

ВИСНОВКИ

У дисертації наведене теоретичне узагальнення і нове вирішення наукової задачі, що виявляється у формалізації біологічних відомостей дендритних обчислень та алокації пам'яті в нейрони зі збільшеною здатністю до активації у моделі просторово-часової асоціативної пам'яті. Також, задача включає проведення аналізу обчислювальних властивостей, можливості практичного застосування, та визначення переваг і недоліків в порівнянні з існуючими підходами.

За підсумками виконаної роботи основні наукові та практичні результати полягають в наступному:

1. Побудована двомірна модель пластичності з різними часовими масштабами сили зв'язку, яка дозволяє навчати стабільні просторові асоціації (зв'язки), виділяти статистично значимі паттерни, та включає сигнал підкріплення. Модель подібна до існуючих підходів, проте включає актуальні біологічні дані взаємодії молекулярних механізмів в синапсах. Робиться припущення про важливість кластеризації АМПА рецепторів для їх стабілізації та підтримки сили зв'язку.

2. Запропонована нова модель асоціативної пам'яті, що використовує нелінійні властивості дендритів та розріджену активацію нейронної мережі, призводить до високої ємності асоціативної пам'яті. Ємність пам'яті на основі сігма-пай нейрона обернено пропорційна розрідженості вихідного вектору та комбінаторно розширеного вхідного вектору. Мультиплікативна активація має кращу синаптичну ємність на 40%, але гіршу (на 400%) інформаційну, згідно порівняння з моделлю Вілшоу.
3. Застосування моделі нейрона з дендритами до задачі прогнозування послідовності показало високу ємність пам'яті: наприклад, для розміру мережі $N=1000$ та 4000 синапсів на нейрон ємність складає 10^5 переходів, порівняно з $\sim 10^4$ для відомої моделі НТМ. Покращення – за рахунок мультиплікативної активації, що дозволяє зменшити кількість зв'язків. Використання часової затримки для контексту спрощує архітектуру мережі, проте накладає обмеження на дальність прогнозування.
4. Побудована модель представлення послідовності в бінарне розріджено-розподілене представлення якісно відтворює такі біологічні ефекти як, збереження схожості, чутливість до порядку, часову схожість. Теоретичні результати підтверджують гіпотезу про важливість алокації пам'яті в нейрони зі збільшеною здатністю до активації для представлення послідовностей. Робота дає стимул для подальшої експериментальної перевірки гіпотези біологами.

Дані моделі застосовувалися в задачі розпізнавання візуальних образів в робототехніці в умовах, коли навчальна вибірка мала, навчання онлайн з динамічною кількістю класів. Проте, якість розпізнавання з використанням біологічно подібної асоціативної пам'яті значно гірша за сучасні методи на основі глибокого навчання. Біологічно подібні алгоритми, що реалізують навчання без учителя ще на зародковому етапі і потребують подальших досліджень.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. **Osaulenko, V.**, Girau, B., Makarenko, O. Henaff, P. Increasing Capacity of Association Memory by Means of Synaptic Clustering. Neural Process. Lett. (2019), pp 1–14 doi:10.1007/s11063-019-10051-7 Особистий внесок: розробка моделі, отримання теоретичних та обчислювальних результатів. (Іноземне видання, реферується в SCOPUS)
2. **Osaulenko V.M.** On the quest to recreate intelligence . Review of Exploring robotic minds : Actions , symbols , and consciousness as self-organizing dynamic. Cognitive Systems Research, 2018, 51, 56–60. <https://doi.org/10.1016/j.cogsys.2018.04.008> (Іноземне видання, реферується в WebOfScience)

3. **Osaulenko V.M** Simple model for sequence prediction based on dendritic spatiotemporal integration System Research & Information Technologies, 2018, № 4, 133-141, DOI:10.20535/SRIT.2308-8893.2018.4.11 (Фахове видання, індексується наукометричними базами DOAJ, Index Copernicus, Google Scholar, PBN)
4. **В.М. Осауленко** Модель інтелектуального агента в одномуірному дискретному світі ISSN 1999-9941, “Інформаційні технології та комп’ютерна інженерія”, 2018, № 3, 30-36, DOI: <https://doi.org/10.31649/1999-9941-2018-43-3-30-36>. (Фахове видання, індексується Google Scholar)
5. **Осауленко В.М.** Двовимірна модель навчання у спайкових нейронних мережах з гомеостазом та навчанням з підкріпленням, System Research & Information Technologies, 2017, № 2, 130-140, DOI: 10.20535/SRIT.2308-8893.2017.1.12 (Фахове видання, індексується наукометричними базами DOAJ, Index Copernicus, Google Scholar, PBN)
6. **Osaulenko V.M.** Testing simple neuron models with dendrites for sparse binary image representation, Штучний інтелект, 2017, № 2, С. 101-108 (Фахове видання, індексується Google Scholar)
7. **Makarenko O.S., Osaulenko V.M.** Application of cellular automates in some models of artificial intelligence. Конференція IEEE SAIC-2018, 2018, 1-4. 10.1109/SAIC.2018.8516837
8. **Osaulenko V.M.** Learning of higher-order correlations improves capacity of an association memory. Nonlinear analysis and applications 4rd International Conference of memory of corresponding member of National Academy of Sciences of Ukraine Valeriy Sergeevich MELNIK (1952-2007) NTUU KPI, Kyiv, Ukraine, April 04-06, 2018, p.76.
9. **Osaulenko V.M.** Model for sequence prediction based on dendritic spatiotemporal integration. 20-th International conference on System Analysis and Information Technology SAIT 2018, May 21-23, 2018, pp.36-37. Institute for Applied System Analysis of National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine
10. **Osaulenko V.M.** Dendritic nonlinearity provides high capacity sequence memory FENS FORUM of neuroscience 2018, p.368, Berlin
11. **Osaulenko V.M.** New ideas to test for artificial neural networks inspired from neuroscience Міжнародна наукова конференція «Сучасна інформатика: проблеми, досягнення та перспективи розвитку» Інституту кібернетики імені В.М.Глушкова, 13-15 грудня 2017, с.82-83
12. **Осауленко В.М.** Тестування простих моделей з дендритами для розрідженого бінарного представлення зображень. XVII Міжнародна науково-технічна конференція Штучний інтелект та інтелектуальні системи. Artificial intelligence and intellegent systems (aiis’2017), Київ 2017, с. 101-108
13. **Осауленко В.М.** На шляху до створення інтелекту Philosophed 2015 міжнародна наукова конференція, ст. 47-49, Київ 2015.

АНОТАЦІЯ

Осауленко В.М. Моделі біологічних нейронних мереж для просторово-часової асоціативної пам’яті. - На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту

Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського” МОН України, Київ, 2019 р.

Дана робота використовує результати дослідження біологічних нейронних мереж для вдосконалення попередніх та побудови нових моделей просторо-часової асоціативної пам'яті. Запропоновано нову модель просторової асоціативної пам'яті на основі сігма-пай нейрона, що враховує дендритні обчислення та має високу ємність пам'яті, що за певними показниками перевершує класичну модель Вілшоу. Показано, що використання нейрона сігма-пай та розрідженої активації покращує на порядок ємність пам'яті переходів послідовності відомої моделі ієрархічно темпоральної пам'яті (НТМ) при однаковій кількості зв'язків. Запропоновано нову модель представлення послідовності в бінарне розріджено-розподілене представлення, яка включає здатність нейрона до активації і якісно відтворює біологічні ефекти, такі як збереження схожості, чутливість до порядку, доповнення послідовності та часову схожість. Представлені моделі просторово-часової асоціативної пам'яті показують гірші результати розпізнавання образів для задач робототехніки в порівнянні з підходом глибоких нейронних мереж, проте вони біологічно подібні, мають привабливі обчислювальні властивості і тому варті подальших досліджень.

Ключові слова: асоціативна пам'ять, дендритні обчислення, прогнозування послідовностей, розпізнавання послідовностей, сігма-пай нейрон, розріджено розподілене представлення

АННОТАЦИЯ

Осауленко В.М. Модели биологических нейронных сетей для пространственно-временной ассоциативной памяти. - На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 - системы и средства искусственного интеллекта. -

Национальный технический университет Украины "Киевский политехнический институт имени Игоря Сикорского" МОН Украины, Киев, 2019.

Данная работа использует результаты исследования биологических нейронных сетей для совершенствования предыдущих и построения новых моделей пространственно-временной ассоциативной памяти. Предложена новая модель пространственной ассоциативной памяти на основе сигма-пай нейрона, что учитывает дендритные исчисления и имеет высокую емкость памяти, которая по определенным показателям превосходит классическую модель Вилшоу. Показано, что использование нейрона сигма-пай и разреженной активации улучшает на порядок емкость памяти переходов последовательности известной модели иерархически темпоральной памяти (НТМ) при одинаковом количестве связей. Предложена новая модель представления последовательности в бинарное разреженно-распределенное представление, что включает способность нейрона к активации и качественно воспроизводит

биологические эффекты, такие как сохранение сходства, чувствительность к порядку, дополнения последовательности и временное сходство. Представленные модели пространственно-временной ассоциативной памяти показывают худшие результаты распознавания образов для задач робототехники в сравнении с подходом глубоких нейронных сетей, однако они биологически подобные, имеют привлекательные вычислительные свойства и интересны для дальнейших исследований.

Ключевые слова: ассоциативная память, дендритные вычисления, прогнозирование последовательностей, распознавание последовательностей, сигма-пай нейрон, разреженно распределенное представление

SUMMARY

Osaulenko V.M. Models of biological neural networks for spatio-temporal association memory. - The manuscript.

Thesis for a candidate degree in specialty 05.13.23 - Systems and applications of artificial intelligence. - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2019.

This work presents biologically plausible models of spatio-temporal association memory based on recent ideas and discoveries from neuroscience. The architectures of most artificial neural networks, like convolutional neural networks, are based on biological ideas dated back to the 60-80s of the previous century. However, since then our understanding of neural computation changed dramatically and new evidence should be formalized, investigated and incorporated into new algorithms.

The work starts with the presentation of a two-dimensional bidirectional plasticity model in a simple spiking neural network. The model includes triplet spike-time-dependent plasticity rule, fast homeostatic plasticity rule and reinforcement signal that significantly speeds up learning and stabilizes the weights. It is based on the findings that molecular mechanisms in the dendritic spines work on different time scales and is similar to eligibility traces. The model shows learning of associative connections that are stable with time, and importantly, new experience does not significantly disrupt previously learned weights. The model has a possible biological interpretation, where the fast weight represents the number and efficiency of AMPA channels and slow weight represents slow molecular mechanisms, possibly CAMKII based, and surface effects on a membrane that determines the stabilization of AMPA receptors. However, it is argued that spiking neural networks are still inapplicable for a practical application like for robotics, so the rest of the work continues with more mathematically abstract, but computationally trackable models.

The next model takes into account dendritic computation, particularly, synaptic clustering. Recently, it was shown that synapses from active input neurons group locally on the dendritic branch, so that later reactivation can cause supralinear summation and dendritic spike. Such coincident detection nature reflected in the sigma-pi neuron model that in this work used to construct association memory. The theoretical and computational analysis shows that multiplicative activation for a sparsely active neuron population gives much higher memory capacity. It reduces the neuron false

activation, known as cross-talk, due to higher-order correlation tracking or coincidence detection in the incoming activity. Comparison to standard Willshaw model shows that the presented model has higher synaptic capacity but lower information capacity, measures that reflect the number of synapses and number of bits per synapse respectively.

The result of higher capacity with sparsely active network and sigma-pi neuron was used for the sequence prediction task. Similarly to time-delay networks, the context was constructed as activity from the last T time steps. The presented model was compared to hierarchical temporal memory (HTM) model that also uses a sparse distributed representation and neuron model with dendrites. The analysis shows higher transition memory capacity by one order of magnitude with the same number of connections. Such difference arises because of the multiplicative neuron sigma-pi model and simplified architecture without minicolumns.

The final model deals with the temporal integration or sequence representation as a sparse distributed binary vector. It is based on the established result that memory is allocated to neurons with increased excitability within hours time scale. This work proposes the hypothesis that excitability is also important on a smaller (seconds) time scale. To check this, the standard algorithm of random projections with k -winners-take-all nonlinearity was modified to include neuronal activity-dependent excitability. The computational simulation shows that the inclusion of excitability creates short term memory and recreates some of the sequence representation effects in biological neural networks. These effects include: similar sequences represented similarly, events in the same sequence represented similarly, beginning of a sequence represented similarly as the whole sequence. Cosine similarity was used to compare different representations. The model can form representation that could be used as a context for the sequence prediction model. Results suggest that the mechanism of memory allocation to excitable neurons is important even for seconds time scale and should be confirmed on biological experiments.

Presented models were investigated for possible practical application for object recognitions in robotics. The task requires learning on a few examples, online, with a non-fixed number of classes; all these restrict from using traditional convolutional neural networks. The natural way to try biologically inspired algorithms for biologically relevant recognition showed poor accuracy results that depend on many handcrafted solutions. The general theory of information processing in biological neural networks is required, however, presented models in this work provide a small contribution towards the main goal.

Keywords: associative memory, dendritic computation, sequence prediction, sequence recognition, sigma-pi neuron, sparse distributed representation